

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2504 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATGGAGGCGA TGCTTCCGCT CTTTGAACCC AAAGGCCGGG TCCTCCTGGT GGACGGCCAC 60
CACCTGGCCT ACCGCACCTT CTTCGCCCTG AAGGGCCTCA CCACGAGCCG GGGCGAACCG 120
GTGCAGGCGG TCTACGGCTT CGCCAAGAGC CTCCTCAAGG CCCTGAAGGA GGACGGGTAC 180
AAGGCCGTCT TCGTGGTCTT TGACGCCAAG GCCCCCTCCT TCCGCCACGA GGCCTACGAG 240
GCCTACAAGG CGGGGAGGGC CCCGACCCCC GAGGACTTCC CCCGGCAGCT CGCCCTCATC 300
AAGGAGCTGG TGGACCTCCT GGGGTTTACC CGCCTCGAGG TCCCCGGCTA CGAGGCGGAC 360
GACGTTCTCG CCACCCTGGC CAAGAAGGCG GAAAAGGAGG GGTACGAGGT GCGCATCCTC 420
ACCGCCGACC GCGACCTCTA CCAACTCGTC TCCGACCGCG TCGCCGTCCT CCACCCCGAG 480
GGCCACCTCA TCACCCCGGA GTGGCTTTGG GAGAAGTACG GCCTCAGGCC GGAGCAGTGG 540
GTGGACTTCC GCGCCCTCGT GGGGGACCCC TCCGACAACC TCCCCGGGGT CAAGGGCATC 600
GGGGAGAAGA CCGCCCTCAA GCTCCTCAAG GAGTGGGGAA GCCTGGAAAA CCTCCTCAAG 660
AACCTGGACC GGGTAAAGCC AGAAAACGTC CGGGAGAAGA TCAAGGCCCA CCTGGAAGAC 720
CTCAGGCTCT CTTTGGAGCT CTCCCGGGTG CGCACCGACC TCCCCCTGGA GGTGGACCTC 780
GCCCAGGGGC GGGAGCCCGA CCGGGAGGGG CTTAGGGCCT TCCTGGAGAG GCTGGAGTTC 840
GGCAGCCTCC TCCACGAGTT CGGCCTCCTG GAGGCCCCCG CCCCCCTGGA GGAGGCCCCC 900
TGGCCCCCGC CGGAAGGGGC CTTCGTGGGC TTCGTCTCTT CCGCCCCCGA GCCCATGTGG 960
GCGGAGCTTA AAGCCCTGGC CGCCTGCAGG GACGGCCGGG TGCACCGGGC AGCAGACCCC 1020
TTGGCGGGGC TAAAGGACCT CAAGGAGGTC CGGGGCCTCC TCGCCAAGGA CCTCGCCGTC 1080
TTGGCCTCGA GGGAGGGGCT AGACCTCGTG CCCGGGGACG ACCCCATGCT CCTCGCCTAC 1140
CTCCTGGACC CCTCCAACAC CACCCCGAG GGGGTGGCGC GGCCTACGG GGGGGAGTGG 1200
ACGGAGGACG CCGCCACCG GGCCCTCCTC TCGGAGAGGC TCCATCGGAA CCTCCTTAAG 1260
CGCCTCGAGG GGGAGGAGAA GCTCCTTTGG CTCTACCACG AGGTGGAAAA GCCCCTCTCC 1320
CGGGTCCTGG CCCACATGGA GGCCACCGGG GTACGGCTGG ACGTGGCCTA CCTTCAGGCC 1380
CTTTCCTTGG AGCTTGCGGA GGAGATCCGC CGCCTCGAGG AGGAGGTCTT CCGCTTGGCG 1440
GGCCACCCCT TCAACCTCAA CTCCCGGGAC CAGCTGGAAA GGGTGCTCTT TGACGAGCTT 1500

```

AGGCTTCCCG CCTTGGGGAA GACGCAAAAG ACAGGCAAGC GCTCCACCAG CGCCGCGGTG 1560
CTGGAGGCCC TACGGGAGGC CCACCCCATC GTGGAGAAGA TCCTCCAGCA CCGGGAGCTC 1620
ACCAAGCTCA AGAACACCTA CGTGGACCCC CTCCCAAGCC TCGTCCACCC GAGGACGGGC 1680
CGCCTCCACA CCCGCTTCAA CCAGACGGCC ACGGCCACGG GGAGGCTTAG TAGCTCCGAC 1740
CCCAACCTGC AGAACATCCC CGTCCGCACC CCCTTGGGCC AGAGGATCCG CCGGGCCTTC 1800
GTGGCCGAGG CGGGTTGGGC GTTGGTGGCC CTGGACTATA GCCAGATAGA GCTCCGCGTC 1860
CTCGCCCACC TCTCCGGGGA CGAAAACCTG ATCAGGGTCT TCCAGGAGGG GAAGGACATC 1920
CACACCCAGA CCGCAAGCTG GATGTTTCGC GTCCCCCGG AGGCCGTGGA CCCCTGATG 1980
CGCCGGGCGG CCAAGACGGT GAACTTCGGC GTCCTCTACG GCATGTCCGC CCATAGGCTC 2040
TCCCAGGAGC TTGCCATCCC CTACGAGGAG GCGGTGGCCT TTATAGAGGC TACTTCCAAA 2100
GCTTCCCCAA GGTGCGGGCC TGGATAGAAA AGACCCTGGA GGAGGGGAGG AAGCGGGGCT 2160
ACGTGGAAAC CCTCTTCGGA AGAAGGCGCT ACGTGCCCGA CCTCAACGCC CGGGTGAAGA 2220
GCGTCAGGGA GGCCGCGGAG CGCATGGCCT TCAACATGCC CGTCCAGGGC ACCGCCGCCG 2280
ACCTCATGAA GCTCGCCATG GTGAAGCTCT TCCCCGCCT CCGGGAGATG GGGGCCCGCA 2340
TGCTCCTCCA GGTCCACGAC GAGCTCCTCC TGGAGGCCCC CCAAGCGCGG GCCGAGGAGG 2400
TGGCGGCTTT GGCCAAGGAG GCCATGGAGA AGGCCTATCC CCTCGCCGTG CCCCTGGAGG 2460
TGGAGGTGGG GATGGGGGAG GACTGGCTTT CCGCCAAGGG TTAG 2504

```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 832 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```

Met Arg Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu
1           5           10           15
Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys Gly
20           25           30
Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala
35           40           45
Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile Val
50           55           60
Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly Gly
65           70           75           80

```

Tyr	Lys	Ala	Gly	Arg 85	Ala	Pro	Thr	Pro	Glu 90	Asp	Phe	Pro	Arg	Gln 95	Leu	
Ala	Leu	Ile	Lys 100	Glu	Leu	Val	Asp	Leu 105	Leu	Gly	Leu	Ala	Arg 110	Leu	Glu	
Val	Pro	Gly 115	Tyr	Glu	Ala	Asp	Asp 120	Val	Leu	Ala	Ser	Leu 125	Ala	Lys	Lys	
Ala	Glu 130	Lys	Glu	Gly	Tyr	Glu 135	Val	Arg	Ile	Leu	Thr 140	Ala	Asp	Lys	Asp	
Leu 145	Tyr	Gln	Leu	Leu	Ser 150	Asp	Arg	Ile	His 155	Val	Leu	His	Pro	Glu	Gly 160	
Tyr	Leu	Ile	Thr	Pro 165	Ala	Trp	Leu	Trp	Glu 170	Lys	Tyr	Gly	Leu	Arg 175	Pro	
Asp	Gln	Trp	Ala 180	Asp	Tyr	Arg	Ala	Leu 185	Thr	Gly	Asp	Glu	Ser 190	Asp	Asn	
Leu	Pro	Gly 195	Val	Lys	Gly	Ile	Gly 200	Glu	Lys	Thr	Ala	Arg 205	Lys	Leu	Leu	
Glu 210	Glu	Trp	Gly	Ser	Leu	Glu 215	Ala	Leu	Leu	Lys	Asn 220	Leu	Asp	Arg	Leu	
Lys 225	Pro	Ala	Ile	Arg	Glu 230	Lys	Ile	Leu	Ala	His 235	Met	Asp	Asp	Leu	Lys 240	
Leu	Ser	Trp	Asp	Leu 245	Ala	Lys	Val	Arg	Thr 250	Asp	Leu	Pro	Leu	Glu 255	Val	
Asp	Phe	Ala	Lys 260	Arg	Arg	Glu	Pro	Asp 265	Arg	Glu	Arg	Leu	Arg 270	Ala	Phe	
Leu	Glu	Arg 275	Leu	Glu	Phe	Gly	Ser 280	Leu	Leu	His	Glu	Phe 285	Gly	Leu	Leu	
Glu	Ser 290	Pro	Lys	Ala	Leu	Glu 295	Glu	Ala	Pro	Trp	Pro 300	Pro	Pro	Pro	Glu	Gly
Ala 305	Phe	Val	Gly	Phe	Val 310	Leu	Ser	Arg	Lys	Glu 315	Pro	Met	Trp	Ala	Asp 320	
Leu	Leu	Ala	Leu	Ala 325	Ala	Ala	Arg	Gly	Gly 330	Arg	Val	His	Arg	Ala 335	Pro	
Glu	Pro	Tyr	Lys 340	Ala	Leu	Arg	Asp	Leu 345	Lys	Glu	Ala	Arg	Gly 350	Leu	Leu	
Ala	Lys	Asp 355	Leu	Ser	Val	Leu	Ala 360	Leu	Arg	Glu	Gly	Leu 365	Gly	Leu	Pro	
Pro	Gly 370	Asp	Asp	Pro	Met	Leu 375	Leu	Ala	Tyr	Leu	Leu 380	Asp	Pro	Ser	Asn	
Thr 385	Thr	Pro	Glu	Gly	Val 390	Ala	Arg	Arg	Tyr	Gly 395	Gly	Glu	Trp	Thr	Glu 400	
Glu	Ala	Gly	Glu	Arg 405	Ala	Ala	Leu	Ser	Glu 410	Arg	Leu	Phe	Ala	Asn 415	Leu	